

A Cycle-Based Bound for Subdominant Eigenvalues of Stochastic Matrices

Steve Kirkland *

Department of Mathematics and Statistics

University of Regina

Regina, Saskatchewan, Canada, S4S 0A2

August 3, 2007

Abstract

Given a primitive stochastic matrix, we provide an upper bound on the moduli of its non-Perron eigenvalues. The bound is given in terms of the weights of the cycles in the directed graph associated with the matrix. The bound is attainable in general, and we characterize a special case of equality when the stochastic matrix has a positive row. Applications to Leslie matrices and to Google-type matrices are also considered.

Keywords: subdominant eigenvalue, stochastic matrix, directed graph

AMS Classifications: 15A42, 15A51

Abbreviated Title: A Cycle-Based Bound for Stochastic Matrices

1 Introduction and preliminaries

A square, entrywise nonnegative matrix S is *stochastic* if each of its row sums is 1 – i.e. if $S\mathbf{1} = \mathbf{1}$, where $\mathbf{1}$ is the all ones vector of the appropriate order. Stochastic matrices arise in the study of discrete time Markov chains that are

*Research partially supported by NSERC under grant number OGP0138251.

time homogenous and have a finite state space. Evidently 1 is an eigenvalue of S , and it follows from Perron-Frobenius theory that for any eigenvalue λ of a stochastic matrix S , we have $|\lambda| \leq 1$. Further, if S is *primitive* – i.e. some power of S has all positive entries, then in fact if $\lambda \neq 1$ is an eigenvalue of S then $|\lambda| < 1$. In the case that S is primitive, the sequence of powers $S^k, k = 1, 2, 3, \dots$ converges as $k \rightarrow \infty$. That convergence is a central component of a corresponding result on the convergence of the Markov chain associated with S . Specifically, when S primitive, the sequence of iterates in the corresponding Markov chain converges to the *stationary vector* for S , i.e. the left Perron vector w^T for S , normalized so that $w^T \mathbf{1} = 1$.

We say that λ is a *subdominant eigenvalue* of a primitive stochastic matrix S if λ is an eigenvalue of S having second-largest modulus after 1. Through a slight abuse of notation, we use $\lambda_2(S)$ to denote a subdominant eigenvalue for S (the abuse arising from the fact that S may have several subdominant eigenvalues). Evidently the rate of convergence of the sequence S^k is governed by the moduli of the subdominant eigenvalues of S . We refer the reader to [8] for further background, terminology and results on stochastic matrices and Markov chains.

For a given stochastic matrix S (or indeed any square entrywise nonnegative matrix), the question of whether S is primitive is purely combinatorial, as it depends only on the positions of the positive entries in S , not on their size. Specifically, let $\Delta(S)$ be the *directed graph* associated with S . That is, if S is $n \times n$, say, the vertices of $\Delta(S)$ are labeled $1, \dots, n$, and there is an arc $i \rightarrow j$ in $\Delta(S)$ if and only if $s_{i,j} > 0$. Then S is primitive if and only if $\Delta(S)$ is strongly connected and, in addition, the greatest common divisor of the lengths of the directed cycles in $\Delta(S)$ is equal to 1. We refer the reader to [1] for these and other results on the relationship between a matrix and its directed graph.

Since the structure of $\Delta(S)$ is influential on the primitivity of S , it is natural to wonder whether the influence of that combinatorial structure can be detected in the moduli of the subdominant eigenvalues of S . In the special case that the Markov chain associated with S is *time reversible*, i.e. $w_i s_{i,j} = w_j s_{j,i}$ for all $i, j = 1, \dots, n$, there is an existing body of work relating the combinatorial features of $\Delta(S)$ to the modulus of $\lambda_2(S)$. The reader is referred to [3] and the references therein for results in that direction. We note in passing that the time reversible setting is highly structured: $\Delta(S)$ can be thought of as an undirected graph (since in the time reversible case, for each $i, j = 1, \dots, n$, $\Delta(S)$ contains the arc $i \rightarrow j$ if and only if it contains

the arc $j \rightarrow i$), while S itself is diagonally similar to a symmetric matrix, and so has all real eigenvalues.

For the case of a general irreducible stochastic matrix S , there are rather fewer results relating $\Delta(S)$ to $\lambda_2(S)$. There is some work in that area, however; for instance, in both [6] and [7], hypotheses on $\Delta(S)$ are used to produce lower bounds on $|\lambda_2(S)|$. The present paper continues in a similar spirit. Specifically, we combine the combinatorial information in the cycles of $\Delta(S)$ with the quantitative information in the entries of S to produce an upper bound on $|\lambda_2(S)|$.

We note that there are a number of bounds on $|\lambda_2(S)|$ in the literature. The following bound, which can be found in [8], will be particularly useful for our purposes.

Proposition 1.1 *Suppose that S is a stochastic matrix of order n . Let*

$$\tau(S) = \max\left\{1 - \sum_{k=1}^n \min\{s_{ik}, s_{jk}\} \mid i, j = 1, \dots, n\right\}.$$

Then for each eigenvalue $\lambda \neq 1$ of S , we have $|\lambda| \leq \tau(S)$.

Observe that $\tau(S) = 1$ if and only if S has a pair of rows with disjoint support - that is, a pair of rows i, j such that $\min\{s_{i,k}, s_{j,k}\} = 0$ for all k . We note in passing that if S is of order 2, then in fact we have $|\lambda_2(S)| = \tau(S)$. However, as the following example shows, the bound provided by Proposition 1.1 may not perform so well, even for matrices of order 3.

Example 1.2 Suppose that $x \in (0, 1)$, and consider the matrix

$$S_x = \begin{bmatrix} x & x(1-x) & (1-x)^2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Then the eigenvalues of S_x are $1, (1-x)e^{\frac{2\pi i}{3}}$ and $(1-x)e^{-\frac{2\pi i}{3}}$, but $\tau(S_x) = 1$ for all $x \in (0, 1)$.

2 A bound on subdominant eigenvalues

Consider a stochastic matrix S , and let C be a cycle in $\Delta(S)$. We use $w(C)$ to denote the product of the entries in S that correspond to the arcs of C .

We begin with a lower bound on the ratio of two entries in the stationary vector for S .

Lemma 2.1 *Suppose that S is an irreducible $n \times n$ stochastic matrix, and that i and j are distinct integers in $\{1, \dots, n\}$. Suppose further that $\Delta(S)$ contains the path $i \equiv k_0 \rightarrow k_1 \rightarrow \dots \rightarrow k_{m+1} \equiv j$. For each $p = 1, \dots, m$ such that $j \rightarrow k_p$, let C_p denote the cycle $k_p \rightarrow \dots \rightarrow k_{m+1} \rightarrow k_p$ in $\Delta(S)$. Then*

$$\frac{w_j}{w_i} \geq \frac{\prod_{p=0}^m s_{k_p, k_{p+1}}}{\prod_{p=1}^{m+1} (1 - s_{k_p, k_p}) - w(C_1) - \sum_{\substack{p=2 \\ j \rightarrow k_p}}^m \prod_{\ell=1}^{p-1} (1 - s_{k_\ell, k_\ell}) w(C_p)}. \quad (2.1)$$

If equality holds in (2.1), then there is just one path from i to j in $\Delta(S)$, namely the path $k_0 \rightarrow k_1 \rightarrow \dots \rightarrow k_{m+1}$.

Proof: We suppose without loss of generality that $j < i$. Let S_i denote the principal submatrix of S formed by deleting row i and column i , and let r^T denote the row vector of order $n - 1$ formed by deleting the i -th entry of $e_i^T S$. Let \hat{w}^T denote the vector formed from w^T by deleting the i -th entry. It now follows from the eigenequation $w^T S = w^T$ that $\hat{w}^T = w_i r^T (I - S_i)^{-1}$, so that $\frac{w_j}{w_i} = r^T (I - S_i)^{-1} e_j$.

Since $\Delta(S)$ contains the path $\equiv k_0 \rightarrow k_1 \rightarrow \dots \rightarrow k_{m+1}$, we have $r^T \geq s_{k_0, k_1} e_{k_1}^T$, while the principal submatrix of S_i on rows and columns k_1, \dots, k_{m+1} dominates (entrywise) the matrix

$$A = \begin{bmatrix} s_{k_1 k_1} & s_{k_1 k_2} & 0 & \dots & 0 & 0 \\ 0 & s_{k_2 k_2} & s_{k_2 k_3} & \dots & 0 & 0 \\ & & \ddots & \ddots & & \vdots \\ 0 & 0 & \dots & 0 & s_{k_m k_m} & s_{k_m k_{m+1}} \\ s_{k_{m+1} k_1} & & \dots & s_{k_{m+1} k_m} & s_{k_{m+1} k_{m+1}} & \end{bmatrix}.$$

Consequently, $r^T (I - S_i)^{-1} e_j \geq s_{k_0, k_1} e_{k_1}^T (I - A)^{-1} e_{k_{m+1}}$, and from the adjoint formula for the inverse of $I - A$, we have

$$s_{k_0, k_1} e_{k_1}^T (I - A)^{-1} e_{k_{m+1}} = \frac{\prod_{p=0}^m s_{k_p, k_{p+1}}}{\prod_{p=1}^{m+1} (1 - s_{k_p, k_p}) - w(C_1) - \sum_{\substack{p=2 \\ j \rightarrow k_p}}^m \prod_{\ell=1}^{p-1} (1 - s_{k_\ell, k_\ell}) w(C_p)}.$$

Thus inequality (2.1) holds.

Note that $(I - S_i)^{-1} = \sum_{q=0}^{\infty} S_i^q$, from which we find that for each $a, b \neq i$, the entry of $(I - S_i)^{-1}$ corresponding to row a column b of S can be written as the sum of the weights of all walks in $\Delta(S)$ that start at vertex a , end at vertex b and that do not go through vertex i (here the weight of a walk is the product of the corresponding entries in S). In particular, if equality holds in (2.1), the only walks in $\Delta(S)$ that start at i , end at j , and do not contain i as an intermediate vertex are the walks already accounted for in the expression for $s_{k_0, k_1} e_{k_1}^T (I - A)^{-1} e_{k_{m+1}}$ above. It now follows that the only path in $\Delta(S)$ from i to j is the path $k_0 \rightarrow k_1 \rightarrow \dots \rightarrow k_{m+1}$. \square

Keeping the notation as above, we make the following definition. For each $i = 1, \dots, n$, we let $\alpha(i, i) = s_{i, i}$, and for any pair of distinct indices $i, j = 1, \dots, n$, such that $j \rightarrow i$ in $\Delta(S)$, we let

$$\alpha(i, j) = \max \left\{ \frac{w(C_0)}{\prod_{p=1}^{m+1} (1 - s_{k_p, k_p}) - w(C_1) - \sum_{\substack{p=2 \\ j \rightarrow k_p}}^m \prod_{\ell=1}^{p-1} (1 - s_{k_\ell, k_\ell}) w(C_p)} \mid i = k_0 \rightarrow \dots \rightarrow k_{m+1} = j \rightarrow k_0 \text{ is a cycle in } \Delta(S) \right\}.$$

We take the convention if $\Delta(S)$ does not contain the arc $j \rightarrow i$, then $\alpha(i, j) = 0$.

Here is one of our main results.

Theorem 2.2 *Suppose that S is an irreducible $n \times n$ stochastic matrix. Then*

$$|\lambda_2(S)| \leq \max \left\{ 1 - \sum_{j=1}^n \min \{ \alpha(i_1, j), \alpha(i_2, j) \} \mid i_1, i_2 = 1, \dots, n \right\}. \quad (2.2)$$

Proof: Denote the stationary vector for S by w^T . Let $W = \text{diag}(w)$ and consider the stochastic matrix T given by $T = W^{-1} S^T W$. For any $i \neq j$ we have $t_{i, j} = \frac{w_j}{w_i} s_{j, i}$, and applying Lemma 2.1, we find that for any cycle $i = k_0 \rightarrow \dots \rightarrow k_{m+1} = j \rightarrow k_0$ in $\Delta(S)$, we have

$$\frac{w_j}{w_i} s_{j, i} \geq \frac{w(C_0)}{\prod_{p=1}^{m+1} (1 - s_{k_p, k_p}) - w(C_1) - \sum_{\substack{p=2 \\ j \rightarrow k_p}}^m \prod_{\ell=1}^{p-1} (1 - s_{k_\ell, k_\ell}) w(C_p)}.$$

It follows that for each $i, j = 1, \dots, n$, $t_{i, j} \geq \alpha(i, j)$. Since T is similar to S^T , we have $|\lambda_2(S)| = |\lambda_2(T)| \leq \max \left\{ 1 - \sum_{j=1}^n \min \{ T_{i_1, j}, T_{i_2, j} \} \mid i_1, i_2 = 1, \dots, n \right\} \leq \max \left\{ 1 - \sum_{j=1}^n \min \{ \alpha(i_1, j), \alpha(i_2, j) \} \mid i_1, i_2 = 1, \dots, n \right\}$. \square

Remark 2.3 Observe that Theorem 2.2 provides a nontrivial bound on $|\lambda_2(S)|$ if and only if no pair of columns in S has disjoint support, or, equivalently, if and only if for each pair of vertices i, j of $\Delta(S)$, there is a vertex k such that $k \rightarrow i$ and $k \rightarrow j$.

Remark 2.4 Note that the cycle structure of $\Delta(S)$ is reflected the quantities $\alpha(i, j)$, and hence in (2.2).

Remark 2.5 The right hand side of (2.2) can be shown to be nondecreasing in each diagonal entry of S .

Example 2.6 Here we revisit the matrix of Example 1.2. Fix $x \in (0, 1)$, and let

$$S_x = \begin{bmatrix} x & x(1-x) & (1-x)^2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Then we have $\alpha(1, 1) = x$, $\alpha(2, 1) = x$ and $\alpha(3, 1) = 1$, while for all other pairs (i, j) , we have $\alpha(i, j) = 0$. From (2.2) we find that $|\lambda_2(S_x)| \leq 1 - x$. As noted in Example 1.2, the eigenvalues of S_x are 1 , $(1-x)e^{\frac{2\pi i}{3}}$ and $(1-x)e^{-\frac{2\pi i}{3}}$. Thus, not only does Theorem 2.2 provide a nontrivial upper bound on $|\lambda_2(S_x)|$, but in fact equality holds in (2.2) for S_x .

Note that for any irreducible stochastic matrix that has a row with all positive entries, the upper bound of Theorem 2.2 is strictly less than 1, by Remark 2.3. Motivated in part by that observation, we next present a bound on $|\lambda_2|$ for irreducible stochastic matrices having a positive row. The proof is immediate from Theorem 2.2.

Corollary 2.7 *Suppose that S is an irreducible stochastic matrix of order n , and that its first row has all positive entries. Let $\beta_1 = s_{1,1}$, and for each $i = 2, \dots, n$, let $\beta_i = \max\left\{\frac{w(C_0)}{\prod_{p=1}^{m+1}(1-s_{k_p, k_p}) - w(C_1) - \sum_{p=2}^m \prod_{\ell=1}^{p-1}(1-s_{k_\ell, k_\ell})w(C_p)} \mid i = k_0 \rightarrow \dots \rightarrow k_{m+1} = 1 \text{ is a path in } \Delta(S), \text{ where } C_p \text{ denotes the cycle } k_p \rightarrow \dots \rightarrow k_{m+1} \rightarrow k_p\right\}$. Finally let $\mu(S) = \min\{\beta_i, i = 1, \dots, n\}$. Then*

$$|\lambda_2(S)| \leq 1 - \mu(S). \tag{2.3}$$

The following is a consequence of Corollary 2.7

Corollary 2.8 *Suppose that S is an irreducible stochastic matrix, and that its first row has all positive entries. Let a be the minimum positive entry in S , and let d be the maximum distance from a vertex $i \in \Delta(S)$ to vertex 1. Then $|\lambda_2(S)| \leq \frac{1-2a+a^{d+2}}{1-2a+a^{d+1}}$.*

Proof: Note that for any cycle C of length m in $\Delta(S)$, $w(C) \geq a^m$. It now follows that if there is a path of length l from vertex $i \neq 1$ to vertex 1 in $\Delta(S)$, then $\beta_i \geq \frac{a^{l+1}}{1-a-a^2-\dots-a^l}$. Since $0 < a \leq \frac{1}{2}$, we find that the expression $\frac{a^{l+1}}{1-a-a^2-\dots-a^l}$ is nonincreasing in l , and that $a \geq \frac{a^{l+1}}{1-a-a^2-\dots-a^l}$. Hence we have $\mu(S) \geq \frac{a^{d+1}}{1-a-a^2-\dots-a^d}$. It now follows from Corollary 2.7 that $|\lambda_2(S)| \leq 1 - \frac{a^{d+1}}{1-a-a^2-\dots-a^d} = \frac{1-2a+a^{d+2}}{1-2a+a^{d+1}}$. \square

3 Characterization of equality in (2.3)

Having established the upper bound (2.2) in Theorem 2.2, and seen via Example 1.2 that equality can hold, it is natural to wonder about which matrices yield equality in the bound. That problem appears to be quite difficult, in part because it requires a characterization of equality in the bound of Proposition 1.1. In this section, we deal with a somewhat less daunting problem: that of characterizing the case of equality in (2.3) of Corollary 2.7.

We begin with four different classes of examples for which equality holds in (2.3).

Example 3.1 Suppose that for some $z \in (0, 1)$, the stochastic matrix S has the form

$$S = \left[\begin{array}{c|c} z & y^T \\ \mathbf{1} & 0 \end{array} \right], \quad (3.4)$$

where I(i) $y_i \geq z(1-z)$, $i = 1, \dots, n-1$, and I(ii) $y^T \mathbf{1} = 1-z$.

Note that $\beta_1 = z$, and that for each $i = 1, \dots, n-1$, $\beta_{i+1} = \frac{y_i}{1-z} \geq z$. Hence we have $\mu(S) = z$, so that by (2.3), $|\lambda_2(S)| \leq 1-z$.

Observe that S has rank 2, and so has just two nonzero eigenvalues. Since S is stochastic, it has 1 as an eigenvalue, and by considering the trace of S , we deduce that $\lambda_2(S) = -(1-z)$. In particular, we see that for the matrix S , equality holds in (2.3).

The set of matrices S of the form (3.4) satisfying conditions I(i) and I(ii) will be referred as *Class I*.

Example 3.2 Suppose that for some $1 \leq p \leq n-2$, S is a stochastic matrix with a positive first row having the form

$$S = \left[\begin{array}{ccccc|ccccc} x_1 & x_2 & \cdots & x_{p-1} & x_p & x_{p+1} & x_{p+2} & \cdots & x_{n-1} & x_n \\ y_2 & 1-y_2 & 0 & \cdots & 0 & 0 & & \cdots & & 0 \\ 0 & y_3 & 1-y_3 & \cdots & 0 & 0 & & \cdots & & 0 \\ \vdots & & \ddots & \ddots & & & & & & \vdots \\ 0 & \cdots & 0 & y_p & 1-y_p & 0 & & \cdots & & 0 \\ \hline 0 & 0 & \cdots & 0 & y_{p+1} & 0 & 0 & \cdots & 0 & 1-y_{p+1} \\ 0 & & \cdots & & 0 & y_{p+2} & 0 & \cdots & 0 & 1-y_{p+2} \\ 0 & & \cdots & & 0 & 0 & y_{p+3} & \cdots & 0 & 1-y_{p+3} \\ \vdots & & & & \vdots & & & \ddots & & \vdots \\ 0 & & \cdots & & 0 & 0 & \cdots & 0 & y_n & 1-y_n \end{array} \right], \quad (3.5)$$

where II(i) $y_i > 0, i = 2, \dots, n, y_{p+1} < 1$, and

II(ii) $m = \gcd\{n-j-p+1 | y_{p+j} < 1\} \geq 2$.

Suppose further that for some $z \in (0, 1)$, the following conditions hold:

II(iii) $x_1 \geq z$;

II(iv) $\frac{y_j x_j}{1 - \sum_{i=1}^{j-1} x_i} \geq z$ for $j = 2, \dots, p$;

II(v) $x_{j+p} y_{p+1} \cdots y_{p+j} = z(1-z)^{j-1} (1 - \sum_{i=1}^p x_i)$ for $j = 1, \dots, n-p$; and

II(vi) $z \sum_{j=1}^{n-p} \frac{(1-z)^{j-1}}{y_{p+1} \cdots y_{p+j}} = 1$.

We note in passing that condition II(vi), in conjunction with condition II(v), ensures that the first row sum for S is 1.

First, we claim that $\mu(S) = z$. To see the claim, first note that $\beta_1 = x_1 \geq z$. For each $j = 2, \dots, p$, we have $\beta_j = \frac{y_2 \cdots y_j x_j}{y_2 \cdots y_{j-1} (1 - \sum_{i=1}^{j-1} x_i)} = \frac{y_j x_j}{1 - \sum_{i=1}^{j-1} x_i}$, which, by condition II(iv), is bounded below by z . Similarly, we find that $\beta_{p+1} = \frac{y_{p+1} x_{p+1}}{1 - \sum_{i=1}^p x_i}$, and that for $j = 2, \dots, n-p$, $\beta_{p+j} = \frac{y_{p+1} \cdots y_{p+j} x_{p+j}}{1 - \sum_{i=1}^p x_i - \sum_{i=1}^{j-1} y_{p+1} \cdots y_{p+i} x_{p+i}}$. It now follows from condition II(v) that $\beta_{p+j} = z$ for $j = 1, \dots, n-p$. In particular, we find that $\mu(S) = z$, and hence $|\lambda_2(S)| \leq 1 - z$.

Next, we claim that equality holds in (2.3). To see this, we first note that the left stationary vector for S , w^T , say, has the property that for each $j = 1, \dots, n$, $\beta_j = \frac{w_1 x_j}{w_j}$, so that in fact equality holds in (2.1) of Lemma 2.1. Letting $T = W^{-1} S^T W$, where $W = \text{diag}(w)$, we see that T can be written as $T = z \mathbf{1} e_1^T + (1-z)M$, where M is a stochastic matrix. Since $\beta_{j+p} = z$ for $j = 1, \dots, n-p$, we find that the principal submatrix of M on rows and columns $p+1, \dots, n$ is stochastic. Further, from condition II(ii), it follows

that that principal submatrix is periodic with period m . Consequently, $e^{\frac{2\pi i}{m}}$ is an eigenvalue of M , so that $(1 - z)e^{\frac{2\pi i}{m}}$ is an eigenvalue of S . Hence $|\lambda_2(S)| = 1 - z$, as claimed.

The set of matrices S of the form (3.5) satisfying II(i)-II(vi) will be referred as *Class II*.

Example 3.3 Suppose that for some $n \geq m \geq 2$ we have $n = rm - 1$ for some integers $r, m \geq 2$, and that S is a stochastic matrix with a positive first row having the form

$$S = \begin{bmatrix} x_1 & x_2 & \dots & x_{n-1} & x_n \\ y_2 & 0 & \dots & 0 & 1 - y_2 \\ 0 & y_3 & 0 & \dots & 1 - y_3 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & 0 & y_n & 1 - y_n \end{bmatrix}, \quad (3.6)$$

where III(i) $y_i > 0, i = 2, \dots, n$, and
III(ii) $y_i < 1$ only if i is a multiple of m .

Suppose also that for some $z \in (0, 1)$, the following conditions hold:

III(iii) $x_i = z(1 - z)^{i-1}, i = 1, \dots, m - 1$;

III(iv) $x_{lm+i} = x_{lm} \left(\frac{1 - \beta_{lm}}{\beta_{lm}} \right) z(1 - z)^{i-1}$ for $i = 1, \dots, m - 1$, and $l = 1, \dots, r - 1$.

It is then straightforward to determine that $\beta_m = \frac{y_m x_m}{(1 - z)^{m-1}}$ and that for $l = 2, \dots, r - 1$,

$$\beta_{lm} = \frac{y_m y_{2m} \dots y_{lm} x_{lm}}{(1 - z)^{m-1} [(1 - z)^{(m-1)(l-1)} - \sum_{j=1}^{l-1} (1 - z)^{(m-1)(l-j-1)} y_m y_{2m} \dots y_{jm} x_{jm}]}.$$

Suppose further that

III(v) $\beta_{lm} \geq z, l = 1, \dots, r - 1$, and

III(vi) $(1 - z)^{m-1} = \sum_{q=1}^{r-1} x_{qm} [(1 - z)^{m-1} + \frac{1 - (1 - z)^{m-1}}{\beta_{lm}}]$.

First, we claim that $\mu(S) = z$. Note that by condition III(v), $\beta_{lm} \geq z$ for $l = 1, \dots, r - 1$. Since $y_i = 1$ for $i = 1, \dots, m - 1$, we have $\beta_j = \frac{x_j}{1 - \sum_{i=1}^{j-1} x_i}$ for $j = 1, \dots, m - 1$, and so condition III(iii), combined with induction on j , shows that $\beta_j = z$ for $j = 1, \dots, m - 1$. Next, fix l between 1 and $r - 1$, and note that $y_{lm+i} = 1$ for $i = 1, \dots, m - 1$. We have $\beta_{lm+i} =$

$$\frac{y_2 \cdots y_{l_{m+i}} x_{l_{m+i}}}{1 - x_1 - \sum_{j=2}^{l_{m+i}-1} y_2 \cdots y_j x_j} = \frac{y_2 \cdots y_{l_m} x_{l_{m+i}}}{1 - x_1 - \sum_{j=2}^{l_m-1} y_2 \cdots y_j x_j - \sum_{j=l_m}^{l_{m+i}-1} y_2 \cdots y_{l_m} x_j} = \frac{x_{l_{m+i}}}{\beta_{l_m} - \sum_{j=l_m}^{l_{m+i}-1} x_j}.$$

Applying condition III(iv), it now follows by induction on i that $\beta_{l_{m+i}} = z$ for $l = 1, \dots, r-1$ and $i = 1, \dots, m-1$. Finally, we note that condition III(vi) simply ensures that when conditions III(iii) and III(iv) hold, the top row of S sums to 1. In particular, it now follows that $\mu(S) = z$, and so by (2.3), we find that $|\lambda_2(S)| \leq 1 - z$.

Next, we claim that equality holds in (2.3). As in Example 3.2, note that the left stationary vector for S , w^T , say, has the property that for each $j = 1, \dots, n$, $\beta_j = \frac{w_1 x_j}{w_j}$, so that in fact equality holds in (2.1) of Lemma 2.1. Setting $W = \text{diag}(w)$ and $T = W^{-1} S^T W$, we see that T can be written as $T = z \mathbf{1} e_1^T + (1-z)M$, where M is a stochastic matrix. Note that $M_{j,1} = \frac{\beta_j - z}{1-z}$ for each $j = 1, \dots, n$. In particular, since $\beta_{l_{m+i}} = z$ for $i = 1, \dots, m-1, l = 0, \dots, r-1$, we have $M_{j,1} > 0$ only if $m|j$; note also that $M_{n,i} > 0$ only if $m|i$ by condition III(ii). It now follows that in $\Delta(M)$, the length of each cycle is divisible by m . We thus find that $e^{\frac{2\pi i}{m}}$ is an eigenvalue of M , so that $(1-z)e^{\frac{2\pi i}{m}}$ is an eigenvalue of S . Hence $|\lambda_2(S)| = 1 - z$, as claimed.

The set of matrices S of the form (3.6) satisfying conditions III(i)-III(vi) will be referred as *Class III*.

Example 3.4 Suppose that for some $n \geq m \geq 2$, and some $a, r \in \mathbb{N}$, we have $n = am - 1 + r$, and that S is a stochastic matrix with a positive first row having the form

$$S = \left[\begin{array}{ccccc|c} x_1 & x_2 & \cdots & x_{n-r-1} & x_{n-r} & \tilde{x}^T \\ y_2 & 0 & \cdots & 0 & v_2 & u_2^T \\ 0 & y_3 & 0 & \cdots & v_3 & u_3^T \\ \vdots & & \ddots & & \vdots & \vdots \\ 0 & \cdots & 0 & y_{n-r} & v_{n-r} & u_{n-r}^T \\ \hline 0 & 0 & \cdots & 0 & \mathbf{1} & 0 \end{array} \right], \quad (3.7)$$

where IV(i) $\tilde{x}^T = [x_{n-r+1} \ \cdots \ x_n]$,

IV(ii) $y_i > 0, i = 2, \dots, n-r$, and

IV(iii) $y_i + v_i + u_i^T \mathbf{1} = 1, i = 2, \dots, n-r$.

Suppose further that IV(iv) $v_i > 0$ only if $m|i$, and that IV(v) $u_i^T \neq 0^T$ only if $m|(i-1)$.

If $a = 1$, suppose also that for some $z \in (0, 1)$,

IV(vi) $x_i = z(1 - z)^{i-1}, i = 1, \dots, m - 1$;

IV(vii) $x_{n-r+j} \geq z(1 - z)^{m-1}, j = 1, \dots, r$; and

IV(viii) $\sum_{j=1}^r x_{n-r+j} = (1 - z)^{m-1}$.

Observe that from the hypotheses on v_i and u_i^T , we find that in fact $y_i = 1$ for $i = 2, \dots, n - r$. Hence we have $\beta_j = \frac{x_j}{1 - \sum_{i=1}^{j-1} x_i}$ for $j = 2, \dots, n - r$, while for $j = n - r + 1, \dots, n$, $\beta_j = \frac{x_j}{1 - \sum_{i=1}^{n-r} x_i}$. Applying conditions IV(vi) and IV(vii), we see that $\beta_j = z, j = 1, \dots, n - r$, while for $j = n - r + 1, \dots, n$, $\beta_j = \frac{x_j}{(1-z)^{m-1}} \geq z$. Thus we have $\mu(S) = z$, so that by (2.3), $|\lambda_2(S)| \leq 1 - z$.

Next, we claim that equality holds in (2.3). As in the previous example, note that equality holds in (2.1) of Lemma 2.1 for the left stationary vector w^T of S . Letting $W = \text{diag}(w)$ and $T = W^{-1}S^T W$, we see that T can be written as $T = z\mathbf{1}e_1^T + (1 - z)M$, where M is a stochastic matrix. From the fact that $\beta_j = z$ for $j = 1, \dots, n - r$, it then follows from conditions IV(iv) and IV(v) that the length of every cycle in $\Delta(M)$ is divisible by m . We thus find that $e^{\frac{2\pi i}{m}}$ is an eigenvalue of M , so that $(1 - z)e^{\frac{2\pi i}{m}}$ is an eigenvalue of S . Hence $|\lambda_2(S)| = 1 - z$, as claimed.

On the other hand, if $a \geq 2$, suppose, in addition to conditions IV(i)-IV(iii), that for some $z \in (0, 1)$,

IV(vi)' $x_i = z(1 - z)^{i-1}, i = 1, \dots, m - 1$;

IV(vii)' $x_{lm+1}y_{lm+1} = z \frac{1 - \beta_{lm}}{\beta_{lm}} x_{lm}$;

IV(viii)' $x_{lm+i} = x_{lm+1}(1 - z)^{i-1}$ for $i = 1, \dots, m - 1$; and

IV(ix)' $x_{n-r+j} \geq x_{n-r}(1 - z), j = 1, \dots, r$.

It is straightforward to show that for each $l = 1, \dots, a - 1$, $\beta_{lm} = \frac{y_2 \dots y_{lm} x_{lm}}{(1-z)^{m-1} [(1-z)^{(m-1)(l-1) - \sum_{j=1}^{l-1} (1-z)^{(m-1)(l-j-1)} y_2 \dots y_{jm} x_{jm}]}$.

Suppose further that

IV(x)' $\beta_{lm} \geq z$ for $l = 1, \dots, a - 1$; and

IV(xi)' $(1 - z)^{m-1} = \sum_{l=1}^{a-1} x_{lm} (1 + (\frac{1 - \beta_{lm}}{\beta_{lm} y_{lm+1}}) (1 - (1 - z)^{m-1})) + \sum_{j=1}^r x_{n-r+j}$.

Note that from the hypotheses on v_i and u_i^T , we find that for each $l = 0, \dots, a - 1$, and $i = 2, \dots, m - 1$, $y_{lm+i} = 1$. Also, for each $j = 2, \dots, n - r$, we have $\beta_j = \frac{y_2 \dots y_j x_j}{1 - x_1 - \sum_{i=2}^{j-1} y_2 \dots y_i x_i}$, while for $j = 1, \dots, r$, we have $\beta_{n-r+j} = \frac{y_2 \dots y_{n-r} x_{n-r+j}}{1 - x_1 - \sum_{i=2}^{n-r} y_2 \dots y_i x_i}$. In particular, for each $i = 1, \dots, n - r - 1$, we find that $\beta_{i+1} = \frac{y_{i+1} x_{i+1} \beta_i}{x_i (1 - \beta_i)}$, while for each $j = 1, \dots, r$, we have $\beta_{n-r+j} = \frac{x_{n-r+j} \beta_{n-r}}{x_{n-r} (1 - \beta_{n-r})}$. From conditions IV(vi)' - IV(viii)', it now follows that for each $j = 1, \dots, n - r$

such that j is not a multiple of m , $\beta_j = z$. Condition IV(ix)' also yields the fact that $\beta_{n-r+j} \geq z, j = 1, \dots, n-r$. From condition IV(x)', we have $\beta_{lm} \geq z, l = 1, \dots, a-1$. Hence we find that $\mu(S) = z$, so that from (2.3), $|\lambda_2(S)| \leq 1-z$.

Next, we claim that equality holds in (2.3). As above, note that the left stationary vector for S , w^T , say, has the property that for each $j = 1, \dots, n-r$, $\beta_j = \frac{w_1 x_j}{w_j}$, while for $j = 1, \dots, r$, $\frac{w_1 x_{n-r+j}}{w_{n-r+j}} \geq \beta_{n-r+j}$. Letting $W = \text{diag}(w)$ and $T = W^{-1}S^T W$, we see that T can be written as $T = z\mathbf{1}e_1^T + (1-z)M$, where M is a stochastic matrix. Note that for $j = 1, \dots, n-r$, $M_{j,1} = \frac{\beta_j - z}{1-z}$, so that in particular, for such j , $M_{j,1} > 0$ only if $m|j$. Note also that for $j = 1, \dots, r$, $M_{n-r+j,i} > 0$ only if $m|(i-1)$, while $M_{n-r,i} > 0$ only if $m|i$.

It now follows that in $\Delta(M)$, the length of each cycle is divisible by m . We thus find that $e^{\frac{2\pi i}{m}}$ is an eigenvalue of M , so that $(1-z)e^{\frac{2\pi i}{m}}$ is an eigenvalue of S . Hence $|\lambda_2(S)| = 1-z$, as claimed.

The set of matrices S of the form (3.7) satisfying conditions IV(i)-IV(v) and either IV(vi)-IV(viii) or IV(vi)'-IV(xi)' will be referred as *Class IV*.

The rest of this section is devoted to providing a converse to Examples 3.1-3.4, namely that the matrices of Classes I-IV are, up to permutation similarity, the only ones yielding equality in (2.3). Throughout the remainder of this section, we assume that \bar{S} is an irreducible stochastic matrix whose first row has all positive entries, and such that equality holds in (2.3). Let w^T be the left stationary vector for \bar{S} , set $W = \text{diag}(w)$, and let $T = W^{-1}\bar{S}^T W$. Finally, let M be given by $M = \frac{1}{1-\mu(\bar{S})}T - \frac{\mu(\bar{S})}{1-\mu(\bar{S})}\mathbf{1}e_1^T$, where $\mu(\bar{S})$ is defined in Corollary 2.3. Observe that M is stochastic; further, since left non-Perron eigenvectors for T are also eigenvectors for M , we find that because $|\lambda_2(T)| = |\lambda_2(\bar{S})| = 1 - \mu(\bar{S})$, it must be the case that $|\lambda_2(M)| = 1$. Throughout, we let $J = \{j \neq 1 | \beta_j = \mu(\bar{S})\}$, and observe that J may be empty.

Lemma 3.5 *Suppose that $i \in J$. Then there is just one path from i to 1 in $\Delta(\bar{S})$, say $i = k_0 \rightarrow \dots \rightarrow k_{m+1} = 1$. Further, each of vertices k_1, \dots, k_{m+1} has indegree two in $\Delta(\bar{S})$.*

Proof: Since $\beta_i = \mu(\bar{S})$, then necessarily equality holds in (2.1); the fact that there is a unique path from i to 1 now follows from Lemma 2.1. Denote that path by $i = k_0 \rightarrow \dots \rightarrow k_{m+1} = 1$. Suppose that for some $j = 1, \dots, m+1$, vertex k_j has indegree at least 3. Then there is a vertex $p \neq 1, k_{j-1}$ such that

$p \rightarrow k_j$. Observe that $p \neq i$, otherwise there is more than one path from i to 1. Considering the walk $k_0 \rightarrow \dots \rightarrow k_{m+1} \rightarrow p \rightarrow k_j \rightarrow \dots \rightarrow k_{m+1}$, and referring to the proof of Lemma 2.1, we see that this walk makes a contribution to $\frac{w_1}{w_i}$ that is unaccounted for in the right side of (2.1); in particular, inequality (2.1) must be strict, a contradiction. We conclude that each of vertices k_1, \dots, k_{m+1} has indegree two, as desired. \square

Corollary 3.6 *Suppose that we have distinct indices $i, j \in J$. Then either the path from i to 1 contains the path from j to 1, or the path from j to 1 contains the path from i to 1.*

Proof: Suppose that the unique paths from i to 1 and j to 1 are $i = a_0 \rightarrow \dots \rightarrow a_p \rightarrow a_{p+1} = 1$ and $j = b_0 \rightarrow \dots \rightarrow b_q \rightarrow b_{q+1} = 1$, respectively. Without loss of generality, we assume that $p \geq q$; we claim that $b_{q+1-k} = a_{p+1-k}$ for $k = 0, 1, \dots, q+1$. We establish the claim by induction on k , and note that certainly the claim holds for $k = 0$, since $b_{q+1} = 1 = a_{p+1}$. Suppose now that for some $k \leq q$ we have $b_{q+1-k} = a_{p+1-k}$. By Lemma 3.5, b_{q+1-k} has indegree two, so that only vertex 1 and vertex b_{q-k} have arcs into b_{q+1-k} . However, since $a_{p+1-k} = b_{q+1-k}$, a similar argument shows that only vertex 1 and vertex a_{p-k} have arcs into a_{p+1-k} . We conclude then that $a_{p-k} = b_{q-k}$, completing the induction step. The claim now follows, and hence we see that the path from i to 1 contains the path from j to 1. \square

Corollary 3.7 *If $J \neq \emptyset$, then there is a unique index $k \in J$ such that for each $j \in J$, the path from k to 1 in $\Delta(\bar{S})$ contains the path from j to 1 in $\Delta(\bar{S})$.*

Proof: For each vertex $i \neq 1$, let $d(i, 1)$ be the distance in $\Delta(\bar{S})$ from i to 1. If J has cardinality one, the result is immediate, so suppose that J has at least two elements. Let k be a vertex in J such that $d(k, 1) \geq d(j, 1)$ for each $j \in J$. Fix a $j \in J$ such that $j \neq k$; by Corollary 3.6, either the path from j to 1 contains the path from k to 1, or vice versa. If the former holds, then $d(j, 1) > d(k, 1)$, a contradiction. It now follows that for each $j \in J$, the path from k to 1 contains the path from j to 1, and that k is the unique vertex in J with that property. \square

Lemma 3.8 *One of the following holds:*

- i) M is reducible with a single periodic essential class;*
- ii) M is irreducible and periodic.*

Proof. Since $|\lambda_2(M)| = 1$, then it is well-known (see [8], for example) that either M has two or more essential classes, or M is reducible with a single periodic essential class, or M is irreducible and periodic. We claim that the first case cannot arise.

To see the claim, let C_1 be the (possibly empty) set of inessential indices for M , and let C_2, \dots, C_k be the classes of essential indices for M (note that $k \geq 3$). It follows that M can be permuted and partitioned as

$$M = \left[\begin{array}{c|c|c|c|c} M_{1,1} & M_{1,2} & M_{1,3} & \dots & M_{1,k} \\ \hline 0 & M_{2,2} & 0 & \dots & 0 \\ \hline 0 & 0 & M_{3,3} & \dots & 0 \\ \hline \vdots & \vdots & & \ddots & \vdots \\ \hline 0 & 0 & \dots & 0 & M_{k,k} \end{array} \right],$$

where the partitioning of M corresponds to the sets C_1, \dots, C_k . Note that if index 1 is not in C_1 , then it follows that \bar{S} is reducible, contrary to hypothesis. Hence $1 \in C_1$. Observe that necessarily, it must be the case that if $i \in C_2$ and $j \in C_3$, we have $\beta_i = \mu(\bar{S}) = \beta_j$. However, it is not the case that in $\Delta(\bar{S})$, either the path from i to 1 contains the path from j to 1, or vice versa, contradicting Corollary 3.6. Hence M has at most one essential class, as claimed. The conclusion now follows. \square

Corollary 3.9 *If $J = \emptyset$, then \bar{S} has the form*

$$\bar{S} = \left[\begin{array}{c|c} \mu(\bar{S}) & y^T \\ \hline \mathbf{1} & 0 \end{array} \right], \quad (3.8)$$

where $y_i > \mu(\bar{S})(1 - \mu(\bar{S}))$, $i = 1, \dots, n-1$ and $y^T \mathbf{1} = 1 - \mu(\bar{S})$. In particular, \bar{S} is in Class I.

Proof. Since $J = \emptyset$, $s_{1,1} = \beta_1 = \mu(\bar{S}) < \beta_i$ for $i = 2, \dots, n-1$. Observe that $\Delta(M)$ differs from $\Delta(\bar{S}^T)$ only in the fact that the former has no loop at vertex 1. In particular, M is irreducible, and, from Lemma 3.8, M is

also periodic. For each $i = 2, \dots, n$, $\Delta(M)$ contains the arc $i \rightarrow 1$, so it follows that the index of periodicity for M is 2. We now deduce that for some positive vector y , \bar{S} has the form

$$\bar{S} = \left[\begin{array}{c|c} \mu(\bar{S}) & y^T \\ \hline \mathbf{1} & 0 \end{array} \right].$$

Evidently it must be the case that $y^T \mathbf{1} = 1 - \mu(\bar{S})$. Further, since $\beta_i = \frac{y_i - 1}{1 - \mu(\bar{S})}$ for $i = 2, \dots, n$, we must have $y_i > \mu(\bar{S})(1 - \mu(\bar{S}))$ for each such i . Thus, \bar{S} is in Class I. \square

Lemma 3.10 *Suppose that M is reducible with a single periodic essential class, say C_2 , and let C_1 denote the set of inessential indices for M . Then $1 \in C_1$, and there is a single pair of vertices u, v such that $u \in C_2, v \in C_1$, and $\Delta(\bar{S})$ contains the arc $u \rightarrow v$. There is a unique path from v to 1 in $\Delta(\bar{S})$. Let k be the unique vertex in J whose path to 1 contains the path from j to 1 for each $j \in J$. Then the subgraph of $\Delta(\bar{S})$ induced by C_2 consists of a single Hamilton path from k to u , along with arcs into k so as to yield a strongly connected periodic directed graph.*

Proof: Observe that M can be permuted and partitioned to the form

$$M = \left[\begin{array}{c|c} M_{1,1} & M_{1,2} \\ \hline 0 & M_{2,2} \end{array} \right],$$

where the partitioning corresponds to C_1 and C_2 . Note that necessarily $1 \in C_1$, otherwise \bar{S} is reducible. Further, we have $C_2 \subseteq J$, and the principal submatrix of M on the rows and columns corresponding to C_2 is irreducible, periodic, and stochastic.

From the fact that the subgraph of $\Delta(\bar{S})$ induced by C_2 is strongly connected, and the fact that each index in C_2 has a unique path to 1, we find that in $\Delta(\bar{S})$ there is a single arc from a vertex in C_2 to a vertex in C_1 , say $u \rightarrow v$. Necessarily, there is a unique path from v to 1 in $\Delta(\bar{S})$. Since the path from k to 1 contains the path from u to 1, we find that necessarily $k \in C_2$ and that the subgraph of $\Delta(\bar{S})$ induced by C_2 contains a Hamilton path from k to u . From Lemma 3.5, each vertex on that Hamilton path that is distinct from k has indegree two in $\Delta(\bar{S})$, so the only possible other arcs in that subgraph are arcs into vertex k . \square

Remark 3.11 It is straightforward to show, using the structure described in Lemma 3.10, that if \bar{S} satisfies the hypotheses of that lemma, then it is permutationally similar to a matrix of the form

$$\left[\begin{array}{ccccc|ccccc} x_1 & x_2 & \dots & x_{p-1} & x_p & x_{p+1} & x_{p+2} & \dots & x_{n-1} & x_n \\ y_2 & 1-y_2 & 0 & \dots & 0 & 0 & & \dots & & 0 \\ 0 & y_3 & 1-y_3 & \dots & 0 & 0 & & \dots & & 0 \\ \vdots & & \ddots & \ddots & & & & & & \vdots \\ 0 & \dots & 0 & y_p & 1-y_p & 0 & & \dots & & 0 \\ \hline 0 & 0 & \dots & 0 & y_{p+1} & 0 & 0 & \dots & 0 & 1-y_{p+1} \\ 0 & & \dots & & 0 & y_{p+2} & 0 & \dots & 0 & 1-y_{p+2} \\ 0 & & \dots & & 0 & 0 & y_{p+3} & \dots & 0 & 1-y_{p+3} \\ \vdots & & & & \vdots & & & \ddots & & \vdots \\ 0 & & \dots & & 0 & 0 & \dots & 0 & y_n & 1-y_n \end{array} \right], \quad (3.9)$$

where a) $y_i > 0, i = 2, \dots, n, y_{p+1} < 1$, and b) $m \equiv \gcd\{n - j - p + 1 | y_{p+j} < 1\} \geq 2$. Here m is the period of the principal submatrix of \bar{S} on the rows and columns corresponding to C_2 .

Lemma 3.12 Suppose that principal submatrix of M on its last $n - p$ rows and columns is irreducible, stochastic, and periodic with period m . Suppose also that \bar{S} has the form (3.9) and satisfies conditions a) and b) of Remark 3.11. Then $x_1 \geq \mu(\bar{S})$, and for $j = 2, \dots, p$ we have $\frac{y_j x_j}{1 - \sum_{i=1}^{j-1} x_i} \geq \mu(\bar{S})$. For $j = 1, \dots, n - p$ we have $x_{j+p} y_{p+1} \dots y_{p+j} = \mu(\bar{S})(1 - \mu(\bar{S}))^{j-1} (1 - \sum_{i=1}^p x_i)$. Finally, we have $\mu(\bar{S}) \sum_{j=1}^{n-p} \frac{(1 - \mu(\bar{S}))^{j-1}}{y_{p+1} \dots y_{p+j}} = 1$. In particular, \bar{S} is in Class II.

Proof: Necessarily we have $\mu(\bar{S}) = \min_j \beta_j$, and from the structure of \bar{S} we see that necessarily $\beta_{p+j} = \mu(\bar{S})$ for $j = 1, \dots, n - p$. It is straightforward to see that for $j = 2, \dots, p, \beta_j = \frac{y_j x_j}{1 - \sum_{i=1}^{j-1} x_i}$, and so the first two conclusions follow from the fact that $\beta_j \geq \mu(\bar{S})$ for $j = 1, \dots, p$.

We find that $\beta_{p+1} = \frac{y_{p+1} x_{p+1}}{1 - \sum_{i=1}^p x_i}$, and that for $j = 2, \dots, n - p, \beta_{p+j} = \frac{y_{p+1} \dots y_{p+j} x_{p+j}}{1 - \sum_{i=1}^p x_i - \sum_{i=1}^{j-1} y_{p+1} \dots y_{p+i} x_{p+i}}$. Since $\beta_{p+j} = \mu(\bar{S}), j = 1, \dots, n - p$, a straightforward induction proof shows that $x_{j+p} y_{p+1} \dots y_{p+j} = \mu(\bar{S})(1 - \mu(\bar{S}))^{j-1} (1 - \sum_{i=1}^p x_i)$ for each such j . The last equation follows from the fact that $\sum_{i=1}^n x_i = 1$. Hence \bar{S} is in Class II. \square

Lemma 3.13 *Suppose that M is irreducible and periodic with period m . Partition the vertex set of $\Delta(M)$ into the classes C_1, C_2, \dots, C_m , so that the only arcs in $\Delta(M)$ are of the form $u \rightarrow v$, where for some $i = 1, \dots, m$, $u \in C_i$ and $v \in C_{(i-1) \bmod m}$. Suppose without loss of generality that $1 \in C_1$. Then $\Delta(M)$ has one of the following two forms:*

i) a Hamilton path from a vertex $k \in C_{m-1}$ to vertex 1, along with possible arcs into k from vertices in C_m and possible arcs from vertex 1 to one or more of the vertices in C_m ;

ii) a collection of vertices $v_1, \dots, v_r \in C_m$ each with a single arc out to a vertex $k \in C_{m-1}$, along with a path from k to 1 through all remaining vertices, possible arcs into k from vertices in C_m , and possible arcs from vertices in C_1 into v_1, \dots, v_r .

Proof: Note that $\cup_{i=1}^{m-1} C_i \subseteq J$, and let k be the vertex in J whose path (in $\Delta(\overline{S})$) to 1 contains all other paths to 1 that start from vertices in J . If it were the case that $k \in C_i$ for some $i = 1, \dots, m-2$, then there would be a vertex $l \in C_{i+1} \subseteq J$ such that $l \rightarrow k$, a contradiction to the fact that the path from k to 1 contains that from l to 1. We conclude that necessarily $k \in C_{m-1}$ or $k \in C_m$.

Suppose that $k \in C_{m-1}$ and consider the path P in $\Delta(\overline{S})$ (and hence in $\Delta(M)$) from k to 1. Observe that each vertex in $\cup_{i=1}^{m-1} C_i$ sits on P . If P is a Hamilton path, then the form described in i) follows readily from the fact that only vertex k can have indegree greater than two in $\Delta(\overline{S})$. If P is not a Hamilton path, then the only vertices not on P are necessarily in C_m . Label these vertices as v_1, \dots, v_r . The form described in ii) now follows upon noting that the only vertices that can have indegree larger than two in $\Delta(\overline{S})$ are k and v_1, \dots, v_r .

Finally, if $k \in C_m$, then in fact the path from k to 1 must be a Hamilton path, otherwise there is a vertex $l \in C_1, l \neq 1$ such that $l \rightarrow k$, a contradiction. The form described in ii) (with $r = 1$ and $v_1 = k$) now follows. \square

Remark 3.14 *Suppose that M satisfies the hypotheses of Lemma 3.13. If form i) holds, then necessarily $n = rm - 1$ for some integer $r \geq 2$. Further, it is straightforward to show that in that case, \overline{S} is permutationally similar*

to a matrix of the form

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_{n-1} & x_n \\ y_2 & 0 & \cdots & 0 & 1 - y_2 \\ 0 & y_3 & 0 & \cdots & 1 - y_3 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & y_n & 1 - y_n \end{bmatrix}, \quad (3.10)$$

where a) $y_i > 0, i = 2, \dots, n$ and b) $y_i < 1$ only if i is a multiple of m .

Lemma 3.15 *Suppose M is irreducible and periodic with period m . Suppose also that \bar{S} has the form (3.10) and satisfies conditions a) and b) of Remark 3.14. Write $n = rm - 1$ for some $r \geq 2$. Then $x_i = \mu(\bar{S})(1 - \mu(\bar{S}))^{i-1}, i = 1, \dots, m-1$, and $x_{lm+i} = x_{lm} \left(\frac{1-\beta_{lm}}{\beta_{lm}} \right) \mu(\bar{S})(1 - \mu(\bar{S}))^{i-1}$ for $i = 1, \dots, m-1$, and $l = 1, \dots, r-1$, where $\beta_{lm} =$*

$$\frac{y_m y_{2m} \cdots y_{lm} x_{lm}}{(1 - \mu(\bar{S}))^{m-1} [(1 - \mu(\bar{S}))^{(m-1)(l-1)} - \sum_{j=1}^{l-1} (1 - \mu(\bar{S}))^{m-1} (l-j-1) y_m y_{2m} \cdots y_{jm} x_{jm}]}$$

for $l = 2, \dots, r-1$. Further, $\beta_m = \frac{y_m x_m}{(1 - \mu(\bar{S}))^{m-1}} \geq \mu(\bar{S})$, and $\beta_{lm} \geq \mu(\bar{S}), l = 1, \dots, r-1$. Finally, we have $(1 - \mu(\bar{S}))^{m-1} = \sum_{q=1}^{r-1} x_{qm} [(1 - \mu(\bar{S}))^{m-1} + \frac{1 - (1 - \mu(\bar{S}))^{m-1}}{\beta_{lm}}]$. In particular, \bar{S} is in Class III.

Proof: Since M is irreducible and periodic with period m , we find that $\beta_{lm+i} = \mu(\bar{S})$ for each $i = 1, \dots, m-1$, and $l = 0, \dots, r-1$. Now $\beta_1 = x_1$, and for each $j \geq 2, \beta_j = \frac{y_2 \cdots y_j x_j}{1 - x_1 - \sum_{i=2}^{j-1} y_2 \cdots y_i x_i}$. Note also that $y_j < 1$ only if m divides j . Since $\beta_j = \mu(\bar{S})$ for $j = 1, \dots, m-1$, an induction proof shows that $x_i = \mu(\bar{S})(1 - \mu(\bar{S}))^{i-1}, i = 1, \dots, m-1$. It now follows that $\beta_m = \frac{y_m x_m}{(1 - \mu(\bar{S}))^{m-1}}$, which must necessarily be bounded below by $\mu(\bar{S})$.

For each $l \geq 1$, we have $\beta_{lm+1} = \frac{y_m \cdots y_{lm} x_{lm+1}}{1 - x_1 - \sum_{i=2}^{lm} y_2 \cdots y_i x_i} = \frac{x_{lm+1}}{\frac{x_{lm}}{\beta_{lm}} - x_{lm}} = \frac{x_{lm+1} \beta_{lm}}{x_{lm} (1 - \beta_{lm})}$.

Since $\beta_{lm+1} = \mu(\bar{S})$, we find that $x_{lm+1} = x_{lm} \mu(\bar{S}) \frac{1 - \beta_{lm}}{\beta_{lm}}$. Further, since $\beta_{lm+i} = \mu(\bar{S})$ for $i = 1, \dots, m-1$, it follows by induction that $x_{lm+i} = x_{lm} \frac{1 - \beta_{lm}}{\beta_{lm}} \mu(\bar{S})(1 - \mu(\bar{S}))^{i-1}, i = 1, \dots, m-1$.

From that we deduce that for each $l \geq 1$,

$$\sum_{i=0}^{m-1} y_2 \cdots y_{lm+i} x_{lm+i} = y_m \cdots y_{lm} x_{lm} [(1 - \mu(\bar{S}))^{m-1} + \frac{1 - (1 - \mu(\bar{S}))^{m-1}}{\beta_{lm}}].$$

An argument by induction on l now shows that

$$\beta_{lm} = \frac{y_m y_{2m} \cdots y_{lm} x_{lm}}{(1 - \mu(\bar{S}))^{m-1} [(1 - \mu(\bar{S}))^{(m-1)(l-1)} - \sum_{j=1}^{l-1} y_m y_{2m} \cdots y_{jm} x_{jm}]}$$

for $l = 2, \dots, r-1$. Evidently $\beta_{lm} \geq \mu(\bar{S})$ for $l = 1, \dots, r-1$.

Finally, the fact that $(1 - \mu(\bar{S}))^{m-1} = \sum_{q=1}^{r-1} x_{qm} [(1 - \mu(\bar{S}))^{m-1} + \frac{1 - (1 - \mu(\bar{S}))^{m-1}}{\beta_{lm}}]$ now follows from the above considerations, and the condition that $\sum_{i=1}^n x_n = 1$. We conclude that \bar{S} is in Class III. \square

Remark 3.16 Suppose that M satisfies the hypotheses of Lemma 3.13. If form ii) holds, then necessarily $n = am - 1 + r$ for some $a, r \in \mathbb{N}$. It is straightforward to show that in that case, \bar{S} is permutationally similar to a matrix of the form

$$\left[\begin{array}{cccccc|c} x_1 & x_2 & \cdots & x_{n-r-1} & x_{n-r} & & \tilde{x}^T \\ y_2 & 0 & \cdots & 0 & v_2 & & u_2^T \\ 0 & y_3 & 0 & \cdots & v_3 & & u_3^T \\ \vdots & & \ddots & & \vdots & & \vdots \\ 0 & \cdots & 0 & y_{n-r} & v_{n-r} & & u_{n-r}^T \\ \hline 0 & 0 & \cdots & 0 & \mathbf{1} & & 0 \end{array} \right], \quad (3.11)$$

where a) $\tilde{x}^T = [x_{n-r+1} \ \cdots \ x_n]$, b) $y_i > 0, i = 2, \dots, n-r$, and c) $y_i + v_i + u_i^T \mathbf{1} = 1, i = 2, \dots, n-r$. Further, d) $v_i > 0$ only if i is a multiple of m , and e) $u_i^T \neq 0^T$ only if $i-1$ is a multiple of m .

The proof of the following result is analogous to that of Lemma 3.15, and is omitted.

Lemma 3.17 *Suppose that M is irreducible and periodic with period m . Suppose also that \bar{S} has the form (3.11), and satisfies conditions a)-e) of Remark 3.16. Write $n-r = am - 1$ for some $a, r \in \mathbb{N}$.*

i) If $n-r = m-1$, then $x_i = \mu(\bar{S})(1 - \mu(\bar{S}))^{i-1}, i = 1, \dots, m-1, x_{n-r+j} \geq \mu(\bar{S})(1 - \mu(\bar{S}))^{m-1}, j = 1, \dots, r$, and $\sum_{j=1}^r x_{n-r+j} = (1 - \mu(\bar{S}))^{m-1}$.

ii) If $n-r = am - 1$ for some $a \geq 2$, then $x_i = \mu(\bar{S})(1 - \mu(\bar{S}))^{i-1}, i = 1, \dots, m-1$. For each $l = 1, \dots, a-1, \beta_{lm} =$

$$\frac{y_2 \cdots y_{lm} x_{lm}}{(1 - \mu(\bar{S}))^{m-1} [(1 - \mu(\bar{S}))^{(m-1)(l-1)} - \sum_{j=1}^{l-1} (1 - \mu(\bar{S}))^{(m-1)(l-j-1)} y_2 \cdots y_{jm} x_{jm}]},$$

$x_{lm+1}y_{lm+1} = \mu(\bar{S}) \frac{1-\beta_{lm}}{\beta_{lm}} x_{lm}$, and $x_{lm+i} = x_{lm+1}(1-\mu(\bar{S}))^{i-1}$ for $i = 1, \dots, m-1$. Also, $x_{n-r+j} \geq x_{n-r}\mu(\bar{S})$, $j = 1, \dots, r$, and finally, we have $(1-\mu(\bar{S}))^{m-1} = \sum_{l=1}^{a-1} x_{lm} \left(1 + \frac{1-\beta_{lm}}{\beta_{lm}y_{lm+1}}\right) (1 - (1-\mu(\bar{S}))^{m-1}) + \sum_{j=1}^r x_{n-r+j}$. In either case, \bar{S} belongs to Class IV.

Here is the main result of this section.

Theorem 3.18 *Suppose that S is an irreducible stochastic matrix whose first row is positive. Equality holds in (2.3) if and only if S is permutationally similar to a matrix in one of Classes I - IV.*

Proof: If S is permutationally similar to a matrix in one of Classes I - IV, then the fact that equality holds in (2.3) follows from Examples 3.1 - 3.4.

Conversely, suppose that equality holds in (2.3) for \bar{S} . By Lemma 3.8, either M is reducible with a single periodic essential class, or M is irreducible and periodic. In the former case, we find from Lemma 3.10, Remark 3.11, and Lemma 3.12 that \bar{S} is permutationally similar to a matrix in Class II.

Suppose next that M is irreducible and periodic. If $J = \emptyset$, then from Corollary 3.9, we find that \bar{S} is in Class I. Now suppose that $J \neq \emptyset$, so that M satisfies Lemma 3.13 i) or ii). If i) holds, then from Remark 3.14 and Lemma 3.15, we find that \bar{S} is permutationally similar to a matrix in Class III. On the other hand, if ii) holds, it follows from Remark 3.16 and Lemma 3.17 that \bar{S} is permutationally similar to a matrix in Class IV. \square

4 Applications and examples

In this section we consider some applications of our results to certain special classes of matrices.

Example 4.1 Suppose that we have positive numbers $x_i, i = 1, \dots, n$, with $\sum_{i=1}^n x_i = 1$. Consider the stochastic companion matrix

$$C = \begin{bmatrix} x_1 & x_2 & \dots & x_{n-1} & x_n \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}.$$

Observe that $\tau(C) = 1$, so that Proposition 1.1 yields no new information on the modulus of $\lambda_2(C)$. However, from Corollary 2.7, we find that $|\lambda_2(C)| \leq 1 - \min\{x_1, \frac{x_i}{1 - \sum_{j=1}^{i-1} x_j} | i = 2, \dots, n\} = \max\{\frac{\sum_{j=i+1}^n x_j}{\sum_{j=i}^n x_j} | i = 1, \dots, n\}$.

Equality can hold in this bound on $|\lambda_2(C)|$. For instance, if $t \in (0, 1)$, and we have the parameters $x_i = t(1-t)^{i-1}, i = 1, \dots, n-1, x_n = (1-t)^{n-1}$, we find that C is in Class IV (with $m = n$ and $a = r = 1$) so that equality holds in (2.3) with $\mu(C) = t$. Indeed when C is constructed with this collection of parameters, then for each n -th root of unity $\omega \neq 1$, we find that $\lambda = \omega(1-t)$ is an eigenvalue of C .

A square nonnegative matrix is a *Leslie matrix* if it has the property that its only positive entries lie in the first row and on the first subdiagonal (and typically that subdiagonal is taken to have all positive entries). Such matrices arise in a discrete-time, age-dependent model of population growth, known as the *Leslie model* (see [5] for a brief description of the model, and [2] for a more comprehensive treatment). In the event that we have a primitive Leslie matrix L , it turns out that the age distributions in the corresponding Leslie model converge to a Perron vector of L . Further, if the Perron value of L is ρ , and λ is an eigenvalue of L of next largest modulus after ρ , then the rate of convergence of the age distributions in the Leslie model is given by $\frac{|\lambda|}{\rho}$. (Note the natural parallel with the behaviour of a Markov chain having a primitive transition matrix.) The next sequence of results applies the ideas of Section 2 to the eigenvalues of certain Leslie matrices.

Theorem 4.2 *Let L be a Leslie matrix of order n , given by*

$$L = \begin{bmatrix} m_1 & m_2 & \dots & m_{n-1} & m_n \\ p_1 & 0 & 0 & \dots & 0 \\ 0 & p_2 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & p_{n-1} & 0 \end{bmatrix},$$

where $m_i > 0, i = 1, \dots, n$ and $p_i > 0, i = 1, \dots, n-1$. Denote the Perron value of L by ρ . Then for any eigenvalue $\lambda \neq \rho$ of L , we have

$$|\lambda| \leq \rho \max\left\{\frac{\sum_{j=i+1}^n p_1 \dots p_{j-1} m_j \rho^{n-j}}{\sum_{j=i}^n p_1 \dots p_{j-1} m_j \rho^{n-j}} | i = 1, \dots, n\right\}. \quad (4.12)$$

Proof: Let $x_1 = \frac{m_1}{\rho}$, and for each $i = 2, \dots, n$, let $x_i = \frac{p_1 \dots p_{i-1} m_i}{\rho^i}$. It is known (see [5]), and not so difficult to show, that $\frac{1}{\rho}L$ is diagonally similar to the matrix

$$\begin{bmatrix} x_1 & x_2 & \dots & x_{n-1} & x_n \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}.$$

Applying the bound of Example 4.1 yields $\frac{\lambda}{\rho} \leq \max\{\frac{\sum_{j=i+1}^n x_j}{\sum_{j=i}^n x_j} | i = 1, \dots, n\}$, which is equivalent to (4.12). \square

Our next result deals with a certain class of stochastic companion matrices.

Theorem 4.3 *Suppose that $2 \leq k \leq \frac{n}{2}$, and that we have positive numbers x_k, \dots, x_n such that $\sum_{j=k}^n x_j = 1$. Consider the companion matrix*

$$S = \begin{bmatrix} 0 & \dots & 0 & x_k & \dots & x_n \\ 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}.$$

Define γ by

$$\gamma = \min\left\{x_k, \dots, x_{2k-1}, \frac{x_{2k} + x_k^2}{\sum_{j=k}^n x_j}, \frac{x_{2k+1} + x_k x_{k+1}}{\sum_{j=k+1}^n x_j}, \dots, \frac{x_n + x_k x_{n-k}}{\sum_{j=n-k}^n x_j}, \frac{x_k x_{n-k+1}}{\sum_{j=n-k+1}^n x_j}, \dots, \frac{x_k x_{n-1}}{\sum_{j=n-1}^n x_j}\right\}.$$

Then $|\lambda_2(S)| \leq (1 - \gamma)^{\frac{1}{k+1}}$.

Proof: It is straightforward to show that

$$e_1^T S^{k+1} e_j = \begin{cases} x_{k+j}, & 1 \leq j \leq k-1 \\ x_{k+j} + x_k x_j, & k \leq j \leq n-k \\ x_k x_j, & n-k+1 \leq j \leq n, \end{cases}$$

so that the first row of S^{k+1} is positive. Letting the left stationary vector of S be w^T , we find from Lemma 2.1 that $\frac{w_1}{w_j} \geq \frac{1}{\sum_{i=j}^n x_i}$ (as it happens, equality

holds). Letting $W = \text{diag}(w)$, we find that the matrix $T = W^{-1}(S^{k+1})^T W$ is stochastic, and its first column is positive. Indeed, $T_{j,1} = \frac{S^{k+1}_{1,j}}{\sum_{i=j}^n x_i}$; the result now follows from the fact that $|\lambda_2(S)| = |\lambda_2(T)|^{\frac{1}{k+1}} \leq \tau(T)^{\frac{1}{k+1}} \leq (1 - \gamma)^{\frac{1}{k+1}}$. \square

We have the following application of Theorem 4.3 to a particular subclass of Leslie matrices.

Corollary 4.4 *Suppose that $2 \leq k \leq \frac{n}{2}$, that $m_i > 0, i = k, \dots, n$, and that $p_i > 0, i = \dots, n - 1$. Let L be the Leslie matrix given by*

$$L = \begin{bmatrix} 0 & \dots & 0 & m_k & \dots & m_n \\ p_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & p_2 & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & p_{n-1} & 0 \end{bmatrix}.$$

Let ρ be the Perron value for L and let

$$\hat{\gamma} = \min \left\{ \frac{p_1 \dots p_{k-1} m_k}{\rho^k}, \dots, \frac{p_1 \dots p_{2k-2} m_{2k-1}}{\rho^{2k-1}}, \frac{p_k \dots p_{2k-1} m_{2k} + p_1 \dots p_{k-1} m_k m_k}{\sum_{i=k}^{2k} p_k \dots p_{i-1} m_i \rho^{2k-i} + \sum_{i=2k+1}^n \frac{p_k \dots p_{i-1} m_i}{\rho^{i-2k}}}, \dots, \frac{p_k \dots p_{n-1} m_n + p_1 \dots p_{n-k-1} m_{n-k} m_k}{\sum_{i=n-k}^n p_k \dots p_{i-1} m_i \rho^{n-i}}, \frac{p_1 \dots p_{n-k} m_{n-k+1} m_k}{\sum_{i=n-k+1}^n p_k \dots p_{i-1} m_i \rho^{n+1-i}}, \dots, \frac{p_1 \dots p_{n-2} m_{n-1} m_k}{\sum_{i=n-1}^n p_k \dots p_{i-1} m_i \rho^{n-1+k-i}} \right\}.$$

Then for any eigenvalue $\lambda \neq \rho$ of L , we have $|\lambda| \leq \rho(1 - \hat{\gamma})^{\frac{1}{k+1}}$.

Proof: For $i = k, \dots, n$, let $x_i = \frac{p_1 \dots p_{i-1} m_i}{\rho^i}$. As in the proof of Theorem 4.2, we find that $\frac{1}{\rho} L$ is diagonally similar to the companion matrix S of Theorem 4.3. The conclusion now follows from Theorem 4.3 after substituting for x_k, \dots, x_n in terms of $p_1, \dots, p_{n-1}, m_k, \dots, m_n$, and then simplifying. \square

We close with an example discussing an application of Theorem 2.2 to a certain iterative method that has been proposed for the computation of Google's PageRank.

Example 4.5 Suppose that we have an $n \times n$ stochastic matrix P , and a positive row vector v^T of order n , normalized so that $v^T \mathbf{1} = 1$. Fix a parameter $c \in (0, 1)$, and consider the stochastic matrix G given by $G = cP + (1-c)\mathbf{1}v^T$. A matrix G of this form is a *Google type* matrix, so named because

a matrix of that type is used in Google's PageRank algorithm, which uses the stationary vector of G in order to rank web pages on the internet. A typical approach to computing the stationary vector for G is via the power method, and it is well known that in the case that P has at least two essential classes (which is certainly the case in practice), or has at least one periodic essential class, the convergence of the power method is geometric, with convergence rate c .

In order to frame our discussion, we impose some assumptions on P . We take P to have essential classes C_1, \dots, C_k , and assume that for any essential index j , $p_{j,j} = 0$ (this happens to be the case in the PageRank application). We suppose also that the rows and columns of P have been simultaneously permuted so that for each $i = 1, \dots, k$, $i \in C_i$. Finally, we assume that the labelling of the rows and columns of P is such that P has the form

$$P = \left[\begin{array}{c|c|c} 0 & P_{12} & P_{13} \\ \hline P_{21} & P_{22} & P_{23} \\ \hline 0 & P_{32} & P_{33} \end{array} \right];$$

here we have partitioned out the first k rows and columns of P , while the second subset of the partition corresponds to those indices i such that $p_{i,j} > 0$ for some $j = 1, \dots, k$, and the third subset of the partition consists of the remaining indices. Partition v^T conformally with P as $v^T = [v_1^T \mid v_2^T \mid v_3^T]$.

In [4], the authors consider an approach to computing the stationary vector of G via a certain iterative aggregation/disaggregation technique, and show that if aggregation is performed on the first k rows and columns, then the convergence is also geometric, and the rate of convergence is given by $c|\lambda_2(H)|$, where H is the following (necessarily irreducible) stochastic matrix:

$$H = \left[\begin{array}{c|c} \frac{P_{22} + cP_{21}P_{12}}{P_{32}} & \frac{P_{23} + cP_{21}P_{13}}{P_{33}} \\ \hline \frac{P_{21}\mathbf{1}}{0} & \left[\begin{array}{c} c\beta v_1^T P_{12} + (1-c)(1+\beta\delta)v_2^T \\ c\beta v_1^T P_{13} + (1-c)(1+\beta\delta)v_3^T \end{array} \right]; \end{array} \right]$$

here $\delta = v_1^T \mathbf{1}$ and $\beta = \frac{1-c}{1-(1-c)\delta}$. Observe that for each index i such that $\Delta(P)$ contains an arc of the form $i \rightarrow j$ for some $j = 1, \dots, k$, the corresponding row of H has all positive entries.

Note that if H is of massive order (as is the case in the PageRank setting), then $|\lambda_2(H)|$ may be difficult to compute, and so an upper bound on $|\lambda_2(H)|$ would then be of some use in estimating the rate of convergence of the iterative method discussed in [4]. Since H has at least one positive row, Theorem 2.2 can be used to bound $|\lambda_2(H)|$.

Let $\sigma = P_{21}\mathbf{1}$, and suppose that σ has order p ; let $u^T = c\beta v_1^T P_{12} + (1 - c)(1 + \beta\delta)v_2^T$ (also of order p) and let $x^T = c\beta v_1^T P_{13} + (1 - c)(1 + \beta\delta)v_3^T$ (of order q , say). Fix an index $j \in \{1, \dots, p\}$. It is straightforward to determine that if $j \neq i \in \{1, \dots, p\}$, then $\alpha(i, j) \geq \sigma_j u_j \frac{\sigma_i u_i}{1 - \sigma_i u_i}$, while $\alpha(j, j) = \sigma_j u_j$. Similarly, if $i \in \{p + 1, \dots, p + q\}$ and $\Delta(P)$ contains the path $i \equiv i_0 \rightarrow i_1 \rightarrow \dots \rightarrow i_{m+1} \equiv j$ (observe that such a path always exists), then

$$\alpha(i, j) \geq \frac{p_{i_0, i_1} \cdots p_{i_m, i_{m+1}} x_{i_0}}{1 - \sigma_j u_j - \sum_{l=1}^m p_{i_l, i_{l+1}} \cdots p_{i_m, i_{m+1}} x_{i_l}}.$$

Thus, these estimates can be used in the bound of Theorem 2.2 to provide an upper bound on $|\lambda_2(H)|$. Consequently, this approach not only provides an alternate proof of the fact (shown in [4]) that $|\lambda_2(H)| < 1$, but also yields quantitative information (in terms of the entries in P , the entries in v^T , and the parameter c) on how close $|\lambda_2(H)|$ can be to 1.

As an illustration of the above, suppose that the minimum positive entry in P is a , and that the minimum positive entry in v^T is b , and that $q \geq 1$. We find that $\delta \geq bk$, that $\beta \geq \frac{1-c}{1-(1-c)bk}$, that $\sigma \geq a\mathbf{1}$, that $u^T \geq \frac{(1-c)b}{1-(1-c)bk}\mathbf{1}^T$, and that $x^T \geq \frac{(1-c)b}{1-(1-c)bk}\mathbf{1}^T$. It now follows that for each $j \in \{1, \dots, p\}$, for the matrix H we have $\alpha(j, j) \geq \frac{(1-c)ab}{1-(1-c)bk}$, while for distinct $i, j \in \{1, \dots, p\}$, then for the matrix H , $\alpha(i, j) \geq \frac{(1-c)^2 a^2 b^2}{(1-(1-c)bk)(1-(1-c)b(k+a))}$. Similarly, if $j \in \{1, \dots, p\}$ and $i \in \{p + 1, \dots, p + q\}$, and if there is a path in $\Delta(P)$ from i to j of length d , then for the matrix H , $\alpha(i, j) \geq \frac{(1-c)(1-a)ba^d}{(1-a)(1-(1-c)b(k+a)) - (1-c)ab(1-a^{d-1})}$. It is not difficult to see that this last quantity, considered as a function of d , is nonincreasing if $1 - a \geq b(1 - c)(k - (k - 2)a - a^2)$, and nondecreasing in d otherwise. Setting $\bar{d}_j = \max_{i \in \{p+1, \dots, p+q\}} d(i, j)$, it follows that for each $j \in \{1, \dots, p\}$ and any $i \in \{1, \dots, p + q\}$, we have for the matrix H that

$$\alpha(i, j) \geq \min\left\{\frac{(1-c)^2 a^2 b^2}{(1-(1-c)bk)(1-(1-c)b(k+a))}, \frac{(1-c)ab}{1-(1-c)bk}, \frac{(1-c)ab}{1-(1-c)b(k+a)}, \frac{(1-c)(1-a)ba^{\bar{d}_j}}{(1-a)(1-(1-c)b(k+a)) - (1-c)ab(1-a^{\bar{d}_j-1})}\right\}.$$

Since $\frac{(1-c)^2 a^2 b^2}{(1-(1-c)bk)(1-(1-c)b(k+a))} \leq \frac{(1-c)ab}{1-(1-c)bk} \leq \frac{(1-c)ab}{1-(1-c)b(k+a)}$, we thus find that $\alpha(i, j) \geq \min\left\{\frac{(1-c)^2 a^2 b^2}{(1-(1-c)bk)(1-(1-c)b(k+a))}, \frac{(1-c)(1-a)ba^{\bar{d}_j}}{(1-a)(1-(1-c)b(k+a)) - (1-c)ab(1-a^{\bar{d}_j-1})}\right\} \equiv \gamma_j$.

Consequently, we have $|\lambda_2(H)| \leq 1 - \sum_{j=1}^k \gamma_j$, thus providing an upper bound on the rate of convergence of the iterative aggregation/disaggregation method discussed in [4]. Observe that this bound, while crude, depends only on the number of essential classes for P , the minimum entries in P and v^T , and upon the lengths of certain paths in $\Delta(P)$.

We note that in certain (highly artificial) examples, in fact the bound of Theorem 2.2 can actually be attained by $|\lambda_2(H)|$. For instance, suppose that $n \geq 3$ and that P is the adjacency matrix of a directed n cycle, say

$$P = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}.$$

In terms of the discussion above, we have $k = 1$, and it turns out that the stochastic matrix H has the form

$$H = \begin{bmatrix} x_1 & x_2 & \dots & x_{n-2} & x_{n-1} \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \text{ where}$$

$$\begin{bmatrix} x_1 & x_2 & \dots & x_{n-2} & x_{n-1} \end{bmatrix} = \frac{1-c}{1-(1-c)v_1} \begin{bmatrix} v_2 & v_3 & \dots & v_n \end{bmatrix} + \frac{c}{1-(1-c)v_1} e_{n-2}^T.$$

Suppose further that for $i = 3, \dots, n-1$, $v_i = v_2 \left(\frac{1-(1-c)(v_1+v_2)}{1-(1-c)v_1} \right)^{i-2}$. Observe that in order for this last condition to hold, it must be the case that $c < \frac{(1-(1-c)(v_1+v_2))^{n-2}}{(1-(1-c)v_1)^{n-3}}$, otherwise $v_n = 1 - \sum_{i=1}^{n-1} v_i \leq 0$ (it is sufficient to take v_2 is close to zero in order to ensure that $1 - \sum_{i=1}^{n-1} v_i > 0$). If all of these conditions hold, it follows that $x_i = \frac{(1-c)v_2}{1-(1-c)v_1} \left(1 - \frac{(1-c)v_2}{1-(1-c)v_1} \right)^{i-1}$, $i = 1, \dots, n-2$. As noted in Example 4.1, equality then holds in (2.3), with $|\lambda_2(H)| = 1 - \frac{(1-c)v_2}{1-(1-c)v_1} = 1 - \mu(H)$.

References

- [1] R. Brualdi and H. Ryser. *Combinatorial Matrix Theory*. Cambridge University Press, Cambridge, 1991.
- [2] H. Caswell. *Matrix Population Models: Construction, Analysis, and Interpretation* (2nd Edition). Sinauer, Sunderland, Massachusetts, 2001.

- [3] P. Diaconis and D. Stroock. Geometric bounds for eigenvalues of Markov chains. *Annals of Applied Probability* 1: 36-61, 1991.
- [4] I. Ipsen and S. Kirkland. Convergence analysis of a PageRank updating algorithm by Langville and Meyer. *SIAM Journal on Matrix Analysis and Applications* 27: 952-967, 2006.
- [5] S. Kirkland. An eigenvalue region for Leslie matrices. *SIAM Journal on Matrix Analysis and Applications*, 13: 507-529, 1992.
- [6] S. Kirkland. A note on the eigenvalues of a primitive matrix with large exponent. *Linear Algebra and its Applications*, 253:103-112, 1997.
- [7] S. Kirkland. Girth and subdominant eigenvalues for stochastic matrices. *Electronic Journal of Linear Algebra*, 12:25-41, 2005.
- [8] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer-Verlag, New York, 1981.